Social Supervision for Parallel Genetic Algorithms

Harold Molina-Bulla and Aníbal R. Figueiras-Vidal

Abstract— **In the present article will be described a first approximation to a system that supervises the genetic training of a set of populations of individuals whose problem to solve is the design of strategies to be confronted in competitions of the vIPD. This supervisor system, called *society*, establishes the parameters of behaviour of the players for their learning process and reproduction process ($p_{learning}$ and $p_{heredity}$ among others) of each of the populations that are training in parallel. The society exchanges information of the behaviour and reproduction parameters and the strategies of the best players every true number of generations of the populations. This exchange of information is carried out in the same way as a Genetic Algorithm, where the individuals of this Meta-population are described by the before mentioned parameters and the strategies of the chosen players. Some conclusions about the Social Supervision as a complementary element for global search and optimisation in parallel genetic algorithms, effects of relative rates for learning and evolution, constructive / destructive interferences, etc., are emphasized; and, finally, suggestions of further research along this direction are presented.**

Index terms—**Genetic Algorithms, Parallel Genetic Algorithms, Baldwin Effect, Lamarck Evolution, Neo-Darwinian Evolution.**

I. Introduction

The Genetic Algorithms as tools of optimisation and global search are widely tested in different theoretical and practical applications. One of the limitations of the Genetic Algorithms based on the Neo-Darwinian Theory consists in not allowing the training in the individuals as method of complementary local search and their poor performance in variable environments.

To solve these limitations the use of additional evolutionary theories such as The Baldwin Effect[1] and the Lamarck Evolution[5] has been posed[7][10][12]. The first allows the learning and evaluate the population after this training, but transfer the original genetic information. The second theory allows the training of the individuals and the transmission of the learned information to their offsprings. These tools, although they improve the services and the speed of convergence of the Genetic Algorithms, present problems of additional energy consumption, propagation of errors made during the training and transmitting information that is not valid due to variations of the environment, when the environment varies in generation in generation[6].

In [8] is planted the use of these evolutionary theories modified to avoid the problems of additional energy consumption and the propagation of errors owed to the learning process of a generation to the following one. These techniques, called Probabilistic Baldwin Effect (PBE) and Probabilistic Lamarckian Evolution (PLE), assign a probability of the learning process being carried out in the individuals ($p_{learning}$), and assign a probability of inheriting the information learned by the parents ($p_{heredity}$). With these techniques besides solving partially the problems posed by the original theories, they allow making use of the advantages of the training and heredity of the information even in variable environments in a certain degree.

On the other side, iterated non-zero sum games are reasonable models for many situations: the best known example being the Iterated Prisoner Dilemma (IPD) which has been considered as representative for the Cold War period and also for pricing policies in a duopolistic market. GA designs have been developed, studied and commented by Axelrod in different papers: [1] is a classical synthesis of these results, which is remarkably interesting from many points of view. The variable Iterated Prisoner Dilemma (vIPD) is a variance of the original game in which the rewards that the player can obtain keep changing throughout the competition, without losing the goal of the game. So, selecting this kind of games to test the PBE and PLE procedures against Neo-Darwinian standard processes seems to be a good way to get representative conclusions.

With this system of Genetic Algorithms and Social Supervision is looked to avoid the exhaustive exploration carried out in [8] of the most adequate parameter for the genetic design of strategies and the training of players. The results are analysed and discussed: showing, in general, an advantage of the complementary evolutionary

Harold Molina-Bulla and Aníbal R. Figueiras-Vidal are with the Signal Theory and Communications Area, *Universidad Carlos III de Madrid*, Spain. E-mail: h.molina@ieee.org.

theories based procedures and social supervision both from the points of view of convergence speed and performance, and more important as the dynamics of the situations are faster.

II. Learning in the Evolution: Baldwin Effect and Lamarck Evolution

A. Evolution Theories and the Genetics Algorithms

J.H. Holland proposed the Genetic Algorithm (GA) based in the Neo-Darwinian evolution theory at the end of the 60. The possible solutions of a problem are individuals with characteristics coded in a binary string (chromosome). The global exploration is realized by combining characteristics extracted from two individuals (the parents) to obtain two offsprings for the next generation with a crossover operator[4]. The selection of the parents is directly related with the fitness of the individuals to solve the problem, the fitness being established by evaluation each string; the individual cannot change its pre-established behaviour. In non-stationary situations, individuals cannot adapt to the changes of the environment, and the population can loose the possibility of getting a reasonable solution.

To solve the adaptation problem we can use others evolution theories, such as the Baldwin Effect, in which the individuals can learn and adapt to the environment their behaviour across their life, achieving a best fitness. In the crossover is used the last evaluation, after the learning an adaptation process, to select the individual whom contributes with their chromosome for the reproduction. The problems of this theory are the costs of the learning process, the errors during this process[6] and the lost of the energy invested in the learning process in each change of generation.

Other evolution theory to use is the Lamarck Evolution, in which the individuals can learn and the learning modifies their chromosomes, allowing transmits the learning to the offsprings. With this theory don not lost the energy invested in the learning process from generation to generation, but if the learning process have errors, this errors are propagated to the next generation and needed more energy to correct them.

B. Heredity vs. Learning Space

To solve the problems of energy invested in the learning and the error propagation between generations we propose in [8] a new space called "Heredity vs. Learning" (see Figure 1) in which we can place a point of probability of learning and transmit the learn to the next generation. Any individual with the point ($p_{Learning}$, $p_{Heredity}$) assigned have a probability $p_{Learning}$ to use the process of learning and have a probability $p_{Heredity}$ to transmit the knowledge acquired with the learning to his offsprings.

To explore and use this space we have proposed to new effects called Probabilistic Baldwin Effect (PBE) and the Probabilistic Lamarck Evolution (PLE).

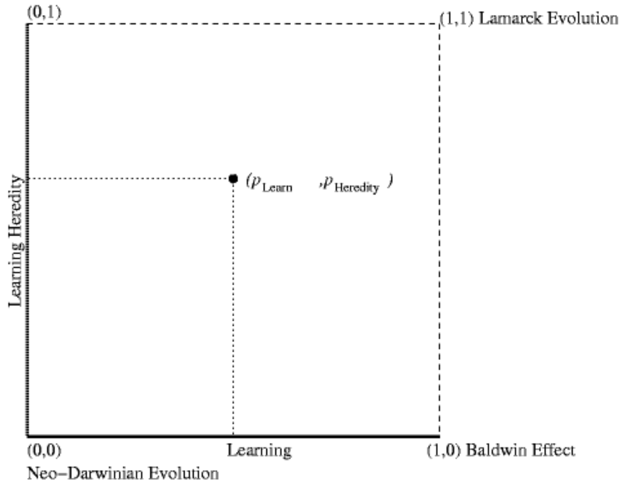

Figure 1 Learning vs. Learning Heredity space: The point p_{Learn}, $p_{Heredity}$ indicate the probability of the individual to learn (p_{Learn}) and the probability of the individual to transmit the learn information to the offsprings ($p_{Heredity}$).

C. Probabilistic Baldwin Effect and Probabilistic Lamarck Evolution

The Probabilistic Baldwin Effect (PBE), proposed in [8] is based on the original Baldwin Effect. The learning process of the individual is conditioned by the probability $p_{Learning}$. Each time the learning process is called, the individual evaluate the probability to apply to this process or not, reducing the energy used in the learning process. This reduces the error introduced by problems with the learning[6].

The Probabilistic Lamarck Evolution (PLE) was proposed in [8] use the $p_{Heredity}$ to select the information to use in the crossover process. The individual has two types of information: the stored in the chromosome, and the learning acquired by the learning process. In the crossover process the individual selects from the first or the second conditioned by the probability $p_{Heredity}$. This allows transmit, in certain degree, the information learned to the offsprings.

The main problem in the use of these tools is to set the correct values of $p_{Learning}$ and $p_{Heredity}$ for the problem to solve. A global exploration can be used, but the computational cost is high and has a low performance.

III. Parallel Genetic Algorithms and Social Supervision

To solve the problems described in [8] we propose the use of an external entity, called Society. This Society establishes the parameters of $p_{Learning}$, $p_{Heredity}$, and other variables related with the learning process, such as the learning cycle longitude.

Under this approximation, we have a set of populations of individuals. Each population is running a Genetic Algorithm and has and independent parameters (such as $p_{Learning}$, $p_{Heredity}$, etc…). These parameters are established by the Society.

The society exchanges information of the behaviour and reproduction parameters and the strategies of the best players in each population every true number of generations and supervise the behaviour of the populations.

The society has the behaviour of a GA, where each population is a meta-individual in the GA of the society; the reproduction parameters, the learning parameters and the chromosomes of the best individuals in the population compose the chromosome of the meta-individual.

This new Meta-GA evolves more slowly, such as the real societies, and has a Lamarck Evolution scheme.

IV. The Iterated Prisoner's Dilemma (IPD)

In 1984, Axelrod[1] proposed a game tournament. In this tournament, the participants send strategies to play a simple game: the Iterated Prisoner's Dilemma. In this game the players can cooperate (C) or defect (D) with the other prisoner at each move. If both cooperate, they are recompensed with R; if one cooperates but the other defects, the first is recompensed with T and the second is punished with S; if both defect, both are punished with P. The relations between the payoffs are $T>R>P>S$.

The Iterated Prisoner's Dilemma is a non-zero sum game, because the payoff of the players can add to a nonzero value. With this situation, both players can obtain more or less payoffs. The rules of the game can favour the cooperation or not: if the payoffs accomplish $2R>T+S$, cooperation emerges. The players do not know the length of the game, because if they know it, they can apply standard optimisation to the last move and go back, and cooperation disappears.

The use of GA in game theory, in particular for the IPD, is very extensive (see [3]); however, the use of GA in non-stationary environments only has been applied when the environment changes from generation to generation [10].

For our Social Supervision experiments, we will create a non-stationary environment where the parameters change slightly for iteration to iteration of IPD, and strongly for generation to generation. Although the environment is non-stationary, i.e. R, T, S, P change, condition $T>R>P>S$ cannot be modified or the target and the game change.

V. Simulations

We establish for the experiments these parameters:

- 20 populations
- 20 generations in the Meta-GA.
- Exchange of information between Meta-Individuals each 2 generations

Each Meta-individual is composed by two populations of vIPD players. The first population evolves with a GA+PLE+PBE and the second population evolves with a GA+Neo Darwinian Evolution (this the reference).

The parameters of each population are:

- 100 Individuals (players)
- 40 generations
- The fitness of the population is the fitness average of the three best individuals obtained after completes the GA training in the population.

The common parameters for the individuals are:

- The chromosome, as defined by Axelrod, uses the outcomes of the two previous moves to determine the current move. The coding for a chromosome was therefore determined by a string of 16 bits, where each bit correspond with the possible instance of the preceding two interactions, and three additional bits that define player's initial moves.

- Mutation probability is 0.015.
- The evaluation method for each member in a population consists in playing IPDs against the three latest best members from the other population. The parameters of each tournament are: initial values of R, T, S, D are 30, 50, 0, 10 respectively; the variation of each parameter is adding up randomly selected points: between –5 to 5 in each move, and among –20 to 20 points between generations; the length of the confrontation is randomly selected between 400 and 450 iterations.
- The learning process for the population with learning capabilities consists on detecting the rule with less average points in a predetermined number of moves, and change the rule. The

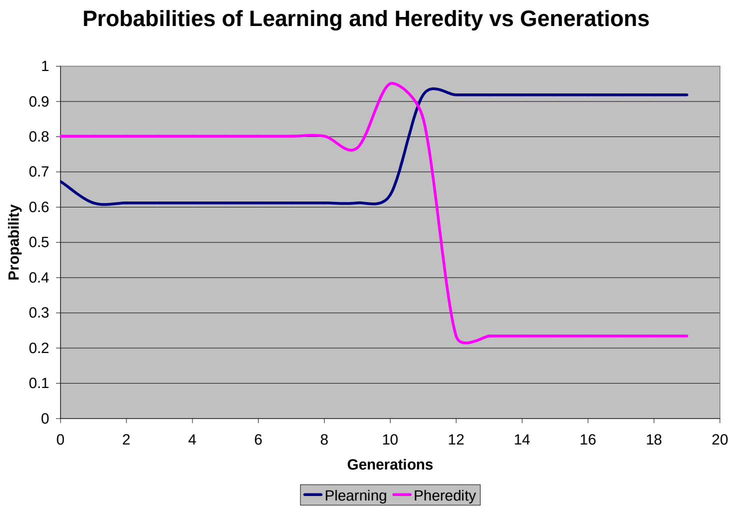

Figure 2 The evolution of the probabilities $p_{Learning}$ and $p_{Heredity}$ of the best population trained with GA+PBE+PLE vs. societies generations.

number of moves, called learning cycle, is established by the society.

We carry out the experiments, stopping at the end of the 20^{th} generation of societies.

In the Figure 2 we can observe the progression of the probabilities of the best individual across the society generations. In this figure we observe the importance of a high value of $p_{Learning}$, and confirm the conclusions obtained in [8], where we say the importance in the use of the learning in the GA in variable environments.

In the Figure 3 we observe the behaviour of the populations trained with GA and PBE and PLE vs. the populations trained with GA and Neo-Darwinian Evolution. The first population obtains better results than the second population after the 9^{th} generation, and achieve a better fitness at the end of the training.

Other results observed after the simulations are the Learning Cycle in the players is in the order of the length of the chromosome, confirming the conclusions in [8]and the importance

VI. Conclusions and Further Works

By including mixed Baldwin Effect and Lamarckian Evolution in GA determination of strategies for the Iterated Prisoner's Dilemma, we have verified that learning is advantageous both for finding solutions in less time and for providing stable solutions in varying environments, assuming that the degree of lamarckism is moderate.

The use of the society to supervise the training of parallel allows find more faster and efficiently the parameters of learning and heredity when we use the Genetic Algorithms with Probabilistic Baldwin Effect and Probabilistic Lamarck Effect. This allows find better strategies with learning to play the vIPD, in comparation with the common GA and Neo-Darwinian Evolution.

The use of parallel GAs increases the computational cost, but reduce the time of execution of the GAs. If we take the advantage of the high degree of parallelism intrinsic in the GAs, and use a cluster of workstastions, the time of execution is reduced.

More detailed experiments are needed, of course: paying attention, in particular, to the following aspects:

- the relationship between the variation rate and the learning parameters;
- the relationship between the length of Meta-GA parameters and the evolution of the populations (Meta-Individuals)
- the dependence with learning parameters in the Meta-GA
- different schemes of exchange information between populations.

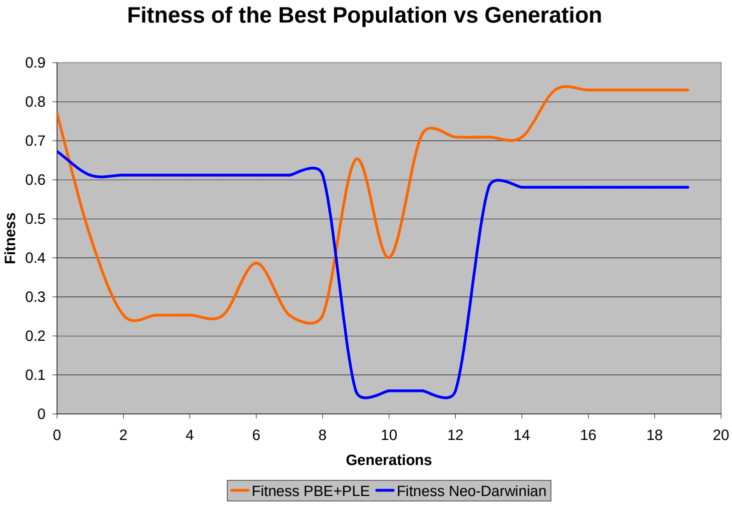

Figure 3 Comparation between the fitness of the best population trained with GA+PBE +PLE and the best population trained with classical GA+ Neo Darwinian in each society generation vs. the societies generations

These questions are addressed by our present work: and we plan, in the future, to consider other games and real examples in strategy design.

VII. Bibliography

[1] R. Axelrod: The Evolution of Cooperation. New York: Basic Books; 1984.

[2] J. M. Baldwin: A New Factor in Evolution; American Naturalist, Vol. 30, pp. 441-451, 1896.

[3] D. F. Fogel: Evolutionary Computation. New York: IEEE Press; 1995.

[4] D. E. Goldberg: Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley; 1989.

[5] J-B M. Lamarck, Philosophie Zoologique, 1809, reedition Lubrecht & Cramer ltd, New York, 1960.

[6] Mayley, G. Landscapes, Learning Costs, and Genetic Assimilation. Evolutionary Computation, Vol. 4 No 3, 1996.

[7] H. Molina-Bulla, A.R. Figueiras-Vidal: The Role Of Learning In Genetic Optimization For Varying Environments: An Assessment Of Baldwin Theory. Proc. Learning 98. Madrid (Spain), 1998.

[8] H. Molina-Bulla: Combinación de Adaptación y Aprendizaje en Optimización de Estrategias (Ph.D. Thesis). Universidad Carlos III de Madrid. Madrid (Spain), 1999.

[9] H. Molina-Bulla,

[10] S. Nolfi, D. Parisi: Learning to Adapt to Changing Environments in Evolving Neural Networks. Adaptive Behavior, Vol. 5, No. 1, pp. 75-98; 1996.

[11] Special Issue on Baldwin Effect, Evolutionary Computation, Vol. 4, No 3, 1996.

[12] P. M. Todd, G. F. Miller: Exploring Adaptive Agency: II. Simulating the Evolution of Associative Learning. In J-A. Meyer, S. W. Wilson (Eds.): From Animals to Animats. Proc. 1st Intl. Conf. on Simulation of Adaptive Behavior. Cambridge, MA: MIT Press; 1991.